

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/575,349  
Source: IFWP  
Date Processed by STIC: 04/25/2006

# ***ENTERED***



IFWP

## RAW SEQUENCE LISTING

DATE: 04/25/2006

PATENT APPLICATION: US/10/575,349

TIME: 08:08:25

Input Set : A:\sequence listing.app

Output Set: N:\CRF4\04252006\J575349.raw

```

3 <110> APPLICANT: Bowles, Dianna et al.
5 <120> TITLE OF INVENTION: Glucosyltransferase
7 <130> FILE REFERENCE: 5585-74167-01
C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/575,349
C--> 9 <141> CURRENT FILING DATE: 2006-04-06
9 <150> PRIOR APPLICATION NUMBER: PCT/GB04/004330
10 <151> PRIOR FILING DATE: 2004-10-12
12 <150> PRIOR APPLICATION NUMBER: 0323813.6
13 <151> PRIOR FILING DATE: 2003-10-13
15 <160> NUMBER OF SEQ ID NOS: 15
17 <170> SOFTWARE: PatentIn version 3.1
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 1464
21 <212> TYPE: DNA
22 <213> ORGANISM: Arabidopsis thaliana
24 <400> SEQUENCE: 1
25 atgaagatta caaaaccaca tgtggccatg ttcgctagcc ccggaatggg ccacatcatc      60
27 ccggtgatcg agctcggaaa acgcttagct gggtcccacg gcttcgatgt caccattttc      120
29 gtccttgaaa ccgacgcagc ctcagctcaa tctcaattcc ttaactcacc aggctgcgac      180
31 gcggcccttg ttgatatcgt tggcctccca acgcccgata tctccggttt agtcgaccca      240
33 tcagcctttt ttgggatcaa gctcttggtc atgatgcgtg agaccattcc taccatccgg      300
35 tcaaagatag aggagatgca acacaaacca acggctctga tcgtagactt gtttggtttg      360
37 gacgcgatac cgctcgggtg tgagttcaac atgttgactt atatcttcat cgcttcaaac      420
39 gcacgttttc tcgcggtggc tttgtttttc ccaacgttgg acaaagacat ggaagaagag      480
41 cacataatca agaagcaacc tatggttatg cctggatgtg aaccggttcg gtttgaagat      540
43 acacttgaaa cattccttga cccaaacagc caactctacc gggaatttgt tcctttcggg      600
45 tcggtttttc caacgtgtga tggattatt gtgaatacat gggatgatat ggagcccaa      660
47 actttgaaat ctcttcaaga cccaaagctc ttgggtcgaa ttgctggtgt accggtttat      720
49 ccaattggtc ctttgtctag accggttgat ccactctaaa ctaatcatcc gggttttgat      780
51 tgggttaaaca aacagccgga cgagtcggtg ctttacattt catttggaag cggtggtctc      840
53 ctctcggcta aacaactaac cgaattggct tggggacttg agatgagtca gcaacgggtc      900
55 gtttggtggg ttcgaccccc ggtggacggt tcagcttgca gtgcatattt atccgctaac      960
57 agtggtaaaa tacgagacgg tacacctgat tatctcccgg aagggtttgt tagccggact      1020
59 catgagagag gctttatggt ctcttcttgg gctcccaaag cggagatctt ggcccaccaa      1080
61 gccgtaggtg ggtttctaac tctctcgggt tggaaatcga ttctcgagag cgtcgttggt      1140
63 ggcgttccga tgatcgctg gccacttttt gcggagcaga tgatgaacgc gacactcctc      1200
65 aacgaagagc ttggcggtgc cgtccgctct aagaaactac cgtcggaggg agtgattacg      1260
67 agggcggaga tcgaggcggt ggtgagaaa atcatggtgg aggaggaagg tgctgagatg      1320
69 agaaagaaga taaagaagct gaaagagacc gctgccgaat cgctgagttg cgacggtgga      1380
71 gtggcgcatg aatcgttgtc aagaatcgcc gacgagagcg agcatctttt ggagcgtgtc      1440
73 aggtgcatgg cacgtggtgc ctag
75                                     1464
76 <210> SEQ ID NO: 2
77 <211> LENGTH: 487

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/575,349

DATE: 04/25/2006

TIME: 08:08:25

Input Set : A:\sequence\_listing.app

Output Set: N:\CRF4\04252006\J575349.raw

```

78 <212> TYPE: PRT
79 <213> ORGANISM: Arabidopsis thaliana
81 <400> SEQUENCE: 2
83 Met Lys Ile Thr Lys Pro His Val Ala Met Phe Ala Ser Pro Gly Met
84 1 5 10 15
87 Gly His Ile Ile Pro Val Ile Glu Leu Gly Lys Arg Leu Ala Gly Ser
88 20 25 30
91 His Gly Phe Asp Val Thr Ile Phe Val Leu Glu Thr Asp Ala Ala Ser
92 35 40 45
95 Ala Gln Ser Gln Phe Leu Asn Ser Pro Gly Cys Asp Ala Ala Leu Val
96 50 55 60
99 Asp Ile Val Gly Leu Pro Thr Pro Asp Ile Ser Gly Leu Val Asp Pro
100 65 70 75 80
103 Ser Ala Phe Phe Gly Ile Lys Leu Leu Val Met Met Arg Glu Thr Ile
104 85 90 95
107 Pro Thr Ile Arg Ser Lys Ile Glu Glu Met Gln His Lys Pro Thr Ala
108 100 105 110
111 Leu Ile Val Asp Leu Phe Gly Leu Asp Ala Ile Pro Leu Gly Gly Glu
112 115 120 125
115 Phe Asn Met Leu Thr Tyr Ile Phe Ile Ala Ser Asn Ala Arg Phe Leu
116 130 135 140
119 Ala Val Ala Leu Phe Phe Pro Thr Leu Asp Lys Asp Met Glu Glu Glu
120 145 150 155 160
123 His Ile Ile Lys Lys Gln Pro Met Val Met Pro Gly Cys Glu Pro Val
124 165 170 175
127 Arg Phe Glu Asp Thr Leu Glu Thr Phe Leu Asp Pro Asn Ser Gln Leu
128 180 185 190
131 Tyr Arg Glu Phe Val Pro Phe Gly Ser Val Phe Pro Thr Cys Asp Gly
132 195 200 205
135 Ile Ile Val Asn Thr Trp Asp Asp Met Glu Pro Lys Thr Leu Lys Ser
136 210 215 220
139 Leu Gln Asp Pro Lys Leu Glu Gly Arg Ile Ala Gly Val Pro Val Tyr
140 225 230 235 240
143 Pro Ile Gly Pro Leu Ser Arg Pro Val Asp Pro Ser Lys Thr Asn His
144 245 250 255
147 Pro Val Leu Asp Trp Leu Asn Lys Gln Pro Asp Glu Ser Val Leu Tyr
148 260 265 270
151 Ile Ser Phe Gly Ser Gly Gly Ser Leu Ser Ala Lys Gln Leu Thr Glu
152 275 280 285
155 Leu Ala Trp Gly Leu Glu Met Ser Gln Gln Arg Phe Val Trp Val Val
156 290 295 300
159 Arg Pro Pro Val Asp Gly Ser Ala Cys Ser Ala Tyr Leu Ser Ala Asn
160 305 310 315 320
163 Ser Gly Lys Ile Arg Asp Gly Thr Pro Asp Tyr Leu Pro Glu Gly Phe
164 325 330 335
167 Val Ser Arg Thr His Glu Arg Gly Phe Met Val Ser Ser Trp Ala Pro
168 340 345 350
171 Gln Ala Glu Ile Leu Ala His Gln Ala Val Gly Gly Phe Leu Thr His
172 355 360 365

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/575,349

DATE: 04/25/2006

TIME: 08:08:25

Input Set : A:\sequence listing.app

Output Set: N:\CRF4\04252006\J575349.raw

```

175 Cys Gly Trp Asn Ser Ile Leu Glu Ser Val Val Gly Gly Val Pro Met
176      370      375      380
179 Ile Ala Trp Pro Leu Phe Ala Glu Gln Met Met Asn Ala Thr Leu Leu
180 385      390      395      400
183 Asn Glu Glu Leu Gly Val Ala Val Arg Ser Lys Lys Leu Pro Ser Glu
184      405      410      415
187 Gly Val Ile Thr Arg Ala Glu Ile Glu Ala Leu Val Arg Lys Ile Met
188      420      425      430
191 Val Glu Glu Glu Gly Ala Glu Met Arg Lys Lys Ile Lys Lys Leu Lys
192      435      440      445
195 Glu Thr Ala Ala Glu Ser Leu Ser Cys Asp Gly Gly Val Ala His Glu
196      450      455      460
199 Ser Leu Ser Arg Ile Ala Asp Glu Ser Glu His Leu Leu Glu Arg Val
200 465      470      475      480
203 Arg Cys Met Ala Arg Gly Ala
204      485
207 <210> SEQ ID NO: 3
208 <211> LENGTH: 1446
209 <212> TYPE: DNA
210 <213> ORGANISM: Arabidopsis thaliana
212 <400> SEQUENCE: 3
213 atgcatatca caaaaccaca cgccgccatg ttttccagtc ccggaatggg ccatgtcatc      60
215 ccggtgatcg agcttggaag gcgtctctcc gctaacaacg gcttccacgt caccgtcttc      120
217 gtcctcgaaa ccgacgcagc ctccgctcaa tccaagttcc taaactcaac cggcgctcgac      180
219 atcgtcaaac ttccatcgcc ggacatttat ggtttagtgg accccgacga ccatgtagtg      240
221 accaagatcg gagtcattat gcgtgcagca gttccagccc tccgatccaa gatcgctgcc      300
223 atgcatcaaa agccaacggc tctgatcggt gacttgtttg gcacagatgc gttatgtctc      360
225 gcaaaggaat ttaacatggt gagttatgtg tttatcccta ccaacgcacg ttttctcgga      420
227 gtttcgattt attatccaaa tttggacaaa gatatcaagg aagagcacac agtgcaaaga      480
229 aaccctctcg ctataccggg gtgtgaaccg gttaggttcg aagatactct ggatgcatat      540
231 ctgggtcccc acgaaccggg gtaccgggat tttgttcgtc atgggtctggc ttacccaaaa      600
233 gccgatggaa ttttggtaaa tacatgggaa gagatggagc ccaaatcatt gaagtccttt      660
235 ctaaacccaa agctcttggg ccgggttgct cgtgtaccgg tctatccaat cggtccttta      720
237 tgcagaccga tacaatcatc cgaaaccgat cacccggttt tggattgggt aaacgaacaa      780
239 ccgaacgagt cggttctcta tatctccttc gggagtgggt gttgtctatc ggcgaaacag      840
241 ttaactgaat tggcgtgggg actcgagcag agccagcaac ggttcgtatg ggtgggtcga      900
243 ccaccggtcg acggttcgtg ttgtagcgag tatgtctcgg ctaacgggtg tggaaccgaa      960
245 gacaacacgc cagagtatct accggaaggg ttcgtgagtc gtactagtga tagaggtttc      1020
247 gtgggtcccc catgggcccc acaagctgaa atcctgtccc atcgggccgt tgggtgggtt      1080
249 ttgaccatt gcggttgagg ctcgacgttg gaaagcgtcg ttggcgccgt tccgatgatc      1140
251 gcatggccac tttttgccga gcagaatatg aatgcggcgt tgctcagcga cgaactggga      1200
253 atcgagtcga gattggatga tccaaaggag gatatttcta ggtggaagat tgaggcgttg      1260
255 gtgaggaagg ttatgactga gaaggaaggt gaagcgatga gaaggaaagt gaagaagttg      1320
257 agagactcgg cggagatgtc actgagcatt gacggtgggt gtttggcgca cgagtcgctt      1380
259 tgcagagtca ccaaggagtg tcaacggttt ttggaacgtg tcgtggactt gtcacgtggt      1440
261 gcttag
264 <210> SEQ ID NO: 4
265 <211> LENGTH: 481
266 <212> TYPE: PRT

```

## RAW SEQUENCE LISTING

DATE: 04/25/2006

PATENT APPLICATION: US/10/575,349

TIME: 08:08:25

Input Set : A:\sequence listing.app

Output Set: N:\CRF4\04252006\J575349.raw

```

267 <213> ORGANISM: Arabidopsis thaliana
269 <400> SEQUENCE: 4
271 Met His Ile Thr Lys Pro His Ala Ala Met Phe Ser Ser Pro Gly Met
272 1 5 10 15
275 Gly His Val Ile Pro Val Ile Glu Leu Gly Lys Arg Leu Ser Ala Asn
276 20 25 30
279 Asn Gly Phe His Val Thr Val Phe Val Leu Glu Thr Asp Ala Ala Ser
280 35 40 45
283 Ala Gln Ser Lys Phe Leu Asn Ser Thr Gly Val Asp Ile Val Lys Leu
284 50 55 60
287 Pro Ser Pro Asp Ile Tyr Gly Leu Val Asp Pro Asp Asp His Val Val
288 65 70 75 80
291 Thr Lys Ile Gly Val Ile Met Arg Ala Ala Val Pro Ala Leu Arg Ser
292 85 90 95
295 Lys Ile Ala Ala Met His Gln Lys Pro Thr Ala Leu Ile Val Asp Leu
296 100 105 110
299 Phe Gly Thr Asp Ala Leu Cys Leu Ala Lys Glu Phe Asn Met Leu Ser
300 115 120 125
303 Tyr Val Phe Ile Pro Thr Asn Ala Arg Phe Leu Gly Val Ser Ile Tyr
304 130 135 140
307 Tyr Pro Asn Leu Asp Lys Asp Ile Lys Glu Glu His Thr Val Gln Arg
308 145 150 155 160
311 Asn Pro Leu Ala Ile Pro Gly Cys Glu Pro Val Arg Phe Glu Asp Thr
312 165 170 175
315 Leu Asp Ala Tyr Leu Val Pro Asp Glu Pro Val Tyr Arg Asp Phe Val
316 180 185 190
319 Arg His Gly Leu Ala Tyr Pro Lys Ala Asp Gly Ile Leu Val Asn Thr
320 195 200 205
323 Trp Glu Glu Met Glu Pro Lys Ser Leu Lys Ser Leu Leu Asn Pro Lys
324 210 215 220
327 Leu Leu Gly Arg Val Ala Arg Val Pro Val Tyr Pro Ile Gly Pro Leu
328 225 230 235 240
331 Cys Arg Pro Ile Gln Ser Ser Glu Thr Asp His Pro Val Leu Asp Trp
332 245 250 255
335 Leu Asn Glu Gln Pro Asn Glu Ser Val Leu Tyr Ile Ser Phe Gly Ser
336 260 265 270
339 Gly Gly Cys Leu Ser Ala Lys Gln Leu Thr Glu Leu Ala Trp Gly Leu
340 275 280 285
343 Glu Gln Ser Gln Gln Arg Phe Val Trp Val Val Arg Pro Pro Val Asp
344 290 295 300
347 Gly Ser Cys Cys Ser Glu Tyr Val Ser Ala Asn Gly Gly Gly Thr Glu
348 305 310 315 320
351 Asp Asn Thr Pro Glu Tyr Leu Pro Glu Gly Phe Val Ser Arg Thr Ser
352 325 330 335
355 Asp Arg Gly Phe Val Val Pro Ser Trp Ala Pro Gln Ala Glu Ile Leu
356 340 345 350
359 Ser His Arg Ala Val Gly Gly Phe Leu Thr His Cys Gly Trp Ser Ser
360 355 360 365
363 Thr Leu Glu Ser Val Val Gly Gly Val Pro Met Ile Ala Trp Pro Leu

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/575,349

DATE: 04/25/2006

TIME: 08:08:25

Input Set : A:\sequence listing.app

Output Set: N:\CRF4\04252006\J575349.raw

```

364      370      375      380
367 Phe Ala Glu Gln Asn Met Asn Ala Ala Leu Leu Ser Asp Glu Leu Gly
368 385      390      395      400
371 Ile Ala Val Arg Leu Asp Asp Pro Lys Glu Asp Ile Ser Arg Trp Lys
372      405      410      415
375 Ile Glu Ala Leu Val Arg Lys Val Met Thr Glu Lys Glu Gly Glu Ala
376      420      425      430
379 Met Arg Arg Lys Val Lys Lys Leu Arg Asp Ser Ala Glu Met Ser Leu
380      435      440      445
383 Ser Ile Asp Gly Gly Gly Leu Ala His Glu Ser Leu Cys Arg Val Thr
384      450      455      460
387 Lys Glu Cys Gln Arg Phe Leu Glu Arg Val Val Asp Leu Ser Arg Gly
388 465      470      475      480
391 Ala
395 <210> SEQ ID NO: 5
396 <211> LENGTH: 1446
397 <212> TYPE: DNA
398 <213> ORGANISM: Arabidopsis thaliana
400 <400> SEQUENCE: 5
401 atgcatatca caaaaccaca cgccgccatg tttccagtc ccggaatggg ccatgtcctc      60
403 ccggtgatcg agctagctaa gcgtctctcc gctaaccacg gcttccacgt caccgtcttc      120
405 gtccttgaaa ctgacgcagc ctccgttcag tccaagctcc ttaactcaac cgggtgttgac      180
407 atcgtcaacc ttccatcgcc cgacatttct ggcttggtag accccaacgc ccatgtggtg      240
409 accaagatcg gagtcattat gcgtgaagct gttccaaccc tccgatccaa gatcgttgcc      300
411 atgcatcaaa acccaacggc tctgatcatt gacttgtttg gcacagatgc gttatgtctt      360
413 gcagcggagt taaacatggt gacttatgtc tttatcgctt ccaacgcgcg ttatctcgga      420
415 gtttcgatat attatccaac tttggacgaa gttatcaaag aagagcacac agtgcaacga      480
417 aaaccgctca ctataccggg gtgtgaaccg gttagatttg aagatattat ggatgcatat      540
419 ctggttcggg acgaaccggg gtaccacgat ttggttcgct actgtctggc ctacccaaaa      600
421 gcggatggaa tcttggtgaa tacatgggaa gagatggagc ccaaatcatt aaagtcctt      660
423 caagaccga aacttttggg ccgggtcgct cgtgtaccgg tttatccggg tgggccgta      720
425 tgcagaccga tacaatcatc cacgaccgat caccgggttt ttgattggtt aaacaaacaa      780
427 ccaaaccgagt cggttctcta catttccttc gggagtgggt gttctctaac ggctcaacag      840
429 ttaaccgaat tggcgtgggg gctcgaggag agccagcaac ggtttatatg ggtggttcga      900
431 ccgcccgttg acggctcgct ttgcagtgat tatttctcgg cttaaaggcgg tgtaacaaaa      960
433 gacaacacgc cagagtatct accagaaggg ttcgtgactc gtacttgcca tagaggtttc      1020
435 atgatcccat catgggcacc gcaagctgaa atcctagccc atcaggccgt tgggtgggtt      1080
437 ttaacacatt gtggttgagg ctcgacgttg gaaagcgtcc tttgcggcgt tccaatgata      1140
439 gcgtggccgc ttttcgccga gcagaatatg aacgcggcgt tgcttagcga tgaactggga      1200
441 atctctgtta gagggtgga tccaaaggag gcgatttcta ggtcgaagat tgaggcgatg      1260
443 gtgaggaagg ttatggctga ggacgaagg gaagagatga gaaggaaagt gaagaagtgt      1320
445 agagacacgg cggagatgtc acttagtatt cacggtgggt gttcggcgca tgagtcgctt      1380
447 tgcagagtca cgaaggagtg tcaacggttt ttggaatgtg tcggggactt gggacgtggt      1440
449 gcttag      1446
452 <210> SEQ ID NO: 6
453 <211> LENGTH: 481
454 <212> TYPE: PRT
455 <213> ORGANISM: Arabidopsis thaliana
457 <400> SEQUENCE: 6

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/575,349

DATE: 04/25/2006

TIME: 08:08:27

Input Set : A:\sequence listing.app

Output Set: N:\CRF4\04252006\J575349.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No

L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date